Are We Compromised? Modelling Security Assessment Games

Viet Pham and Carlos Cid

Information Security Group
Royal Holloway, University of London
Egham, Surrey TW20 0EX, United Kingdom
{viet.pham.2010,carlos.cid}@rhul.ac.uk

Abstract. Security assessments are an integral part of organisations' strategies for protecting their digital assets and critical IT infrastructure. In this paper we propose a game-theoretic modelling of a particular form of security assessment – one which addresses the question "are we compromised?". We do so by extending the recently proposed game "FlipIt", which itself can be used to model the interaction between defenders and attackers under the Advanced Persistent Threat (APT) scenario. Our extension gives players the option to "test" the state of the game before making a move. This allows one to study the scenario in which organisations have the option to perform periodic security assessments of such nature, and the benefits they may bring.

1 Introduction

The protection of digital assets and critical IT infrastructure is an ever-growing concern for individuals, companies and nations. Information security is now a priority area for investment, given the growing threats from hackers, competitors, organised criminal gangs and enemy nation-states, and the potential for loss of privacy and revenue, negative reputational impact and effects in public welfare. In addition to direct investment in suitable and robust IT infrastructure, the performance of frequent security assessments is also considered an important component of the defense strategy against cyber-attacks. A security assessment is the process of determining how effectively an entity being assessed meets specific security objectives [11]. A common method of assessment is a *penetration testing*, where security professionals target the network and other IT resources, to try to identify and verify any vulnerabilities found. Popular penetration testing methodologies and frameworks work by essentially *mimicking* the popular forms of attack used by hackers.

The nature of cyber attacks has however been steadily changing in recent years. While previously the typical threats were *script kiddies*, more interested in defacing websites for fun and pride, attacks motivated by financial gains are increasingly becoming more prevalent. Particularly in the corporate and government spheres, the threat of espionage and theft of intellectual property and state secrets are growing causes of concern. With these goals in mind, the methods

used by attackers have also evolved. A form of attack that has received much attention recently are the so-called *Advanced Persistent Threats* (APT), which can often be seen as a signal of international cyber warfare [2]. The premises in this form of attack are that IT networks and systems *are* vulnerable, and therefore can be compromised by adversaries with enough resources and motivation; furthermore, attacks are stealthy in nature [15, 5], and adversaries can remain in control of the network and systems for long periods without detection. Recent examples of cyber attacks that fit this profile are the security breach at RSA Data Security [6], and the Stuxnet [9] worm infection of Iranian systems.

These developments should in turn motivate a reflection on whether current methods of security assessment remain sound under the changing nature of attacks. A security assessment is typically seen to be trying to answer the question "are we vulnerable?" (and if so, how can we fix it?). Under APT's premise, the answer for this question is certainly "yes". Thus a security assessment needs also to address the question "are we compromised?", and organisations need to consider cost-effective ways in which they can can regain control of their IT assets if the answer is positive. This current gap should certainly be the cause of concern for professionals involved in the security of highly-targeted organisations.[1]

In this paper we propose a simple game-theoretic modelling of this form of security assessment, and study its application in 2-player security games. Game modelling has been shown to be useful in studying strategic decisions toward a wide range of security problems, from technical [10] to managerial [8, 13]. Our model extends the recently proposed game "FlipIt" [14], which itself can be used to model the interaction between defenders and attackers under the APT scenario. Our extension gives players the option to "test" the state of the game (i.e. answer the question "are we compromised"). This allows one to study the scenario in which organisations have the option of performing periodic security assessments of such nature, and the benefits they may bring. In particular, how these assessments can fit into an organisation's security investment strategy. Proposals of models for security investment and security testing have appeared before in the literature (e.g. [7, 4, 3]); here we leverage on the elegance of FlipIt to investigate strategies for the application of this form of security assessment.

This paper is organised as follows. In Section 2 we describe the game "FlipIt". In Section 3 we propose our extension to the game, by introducing the option of a security assessment which discloses the state of the game. We study further extensions in Sections 4 and 5. We finish with our conclusions in Section 6.

2 FlipIt: the game

The original FlipIt games [14] capture the battle between a defender and an advanced persistent threat (APT) attacker for the control of a resource. The

[1] In fact these points were emphatically argued in a recent testimony before the U.S.-China Economic and Security Review Commission Hearing on "Developments in China's Cyber and Nuclear Capabilities", where one of the participants stressed the need of periodic security assessments of the latter nature [1].

game is modelled over infinite time, in which a player makes a move to gain control of the resource; it remains in this state until the opponent makes its own move to take over. This control-alternating process repeats infinitely as time passes, and the utility of each player is determined by the total/average amount of time it controls the resource, as well as the cost required to take over the resource from its opponent.

Formally, the defender and the attacker are denoted by player 0 and 1, respectively. The game timeline starts from some moment $t = 0$ and is continuously indefinite, so that the amount of control time for each player can be conveniently computed in $\mathbb{R}$. Let $C_i(t)$ be 1 if player i controls the resource at time t, and 0 otherwise. For example, if the defender moves at time t, then $C_0(t) = 1$; similarly, we have $C_0(t') = 0$ if the attacker moves at time t'. This allows the total control time of player i until time t to be computed as

$$G_i(t) = \int_0^t C_i(t)dt.$$

Denote player i's number of moves until time t by $n_i(t)$, and the constant cost for each move by k_i; then the *net benefit* of player i is given by

$$B_i(t) = G_i(t) - n_i(t)k_i.$$

Alternatively, since the game continues indefinitely, a player's utility can be represented by its *average benefit* per unit time:

$$\beta_i(t) = \frac{B_i(t)}{t} = \frac{G_i(t)}{t} - \frac{n_i(t)}{t}k_i = \gamma_i(t) - \alpha_i(t)k_i.$$

We call $\gamma_i(t)$ and $\alpha_i(t)$ the *average gain rate* and the *average move rate* of player i up to time t, respectively. One may further assume that the functions $\gamma_i(t)$ and $\alpha_i(t)$ converge to the values γ_i and α_i, respectively, as $t \to \infty$. We can then conveniently represent player i's utility without the time dimension as simply

$$\beta_i = \lim_{t\to\infty} \beta_i(t) = \gamma_i - \alpha_i k_i.$$

What remains to be modelled are γ_i and α_i, which strongly depend on how the players strategically act in the game. While the authors in [14] discuss several types of strategies for each player, in this paper we focus only on the so-called *periodic strategies with random phase*, which is the main tool for our work. In periodic games, we assume that before start, each player chooses a rate $\alpha_i > 0$ so that as the game progresses, player i moves at rate α_i, i.e., after every $\delta_i = 1/\alpha_i$ units of time. Furthermore, player i does not start moving immediately at $t = 0$, but selects uniformly at random a starting point in the interval $[0, \delta_i]$; this is called *phase*. While player i cannot control its phase, its game action is determined by the chosen move rate α_i. For convenience, we denote the action space for periodic moving strategies for both players as

$$P = \{P_\alpha | \alpha > 0\}.$$

Since players move periodically, their expected average control time γ_i, or average gain, can be computed in the following two cases:

- $\alpha_0 \geq \alpha_1$: let $r = \alpha_1/\alpha_0 = \delta_0/\delta_1$; we note that for every attacker's period interval $[t^*, t^* + \delta_1]$, the defender moves at time t uniformly random within $[t^*, t^* + \delta_0]$, yielding a gain $t^* + \delta_1 - t$, which can be expectedly computed as

$$G_0^* = \int_{t^*}^{t^*+\delta_0} \frac{t^* + \delta_1 - t}{\delta_0} dt = \delta_1 - \frac{\delta_0}{2} = \delta_1(1 - \frac{r}{2}).$$

  This implies that the defender's average gain is $\gamma_0 = G_0^*/\delta_1 = 1 - r/2$; it also means that the attacker's average gain is $\gamma_1 = 1 - \gamma_0 = r/2$. Therefore, we have the players' utilities as

$$\beta_0(\alpha_0, \alpha_1) = 1 - \frac{r}{2} - \alpha_0 k_0 = 1 - \frac{\alpha_1}{2\alpha_0} - \alpha_0 k_0,$$
$$\beta_1(\alpha_0, \alpha_1) = \frac{r}{2} - \alpha_1 k_0 = \frac{\alpha_1}{2\alpha_0} - \alpha_1 k_1.$$

- $\alpha_0 \leq \alpha_1$: similar analysis gives the following

$$\beta_0(\alpha_0, \alpha_1) = \frac{r}{2} - \alpha_0 k_0 = 1 - \frac{\alpha_0}{2\alpha_1} - \alpha_0 k_0,$$
$$\beta_1(\alpha_0, \alpha_1) = 1 - \frac{r}{2} - \alpha_1 k_0 = \frac{\alpha_0}{2\alpha_1} - \alpha_1 k_1.$$

We note that when a player has lost the control due to the opponent's move, it does not immediately move to regain it but rather needs to wait for its periodic move. This is because moves are presumably "stealthy", and neither player knows at any time who is controlling the resource. In addition to the periodic move scenario, [14] also studies strategies involving randomised moves, as well as adaptive strategies based on the opponent's past moves. Although we do not consider these here, we note that the modelling presented in this paper may be similarly applied to other scenarios discussed in [14].

The main reason for choosing FlipIt to base our work on is its simple, though elegant, modelling of real-world IT security defender-attacker interaction. Indeed, strategies for organisational security are often determined in the very early phase of the business, and they are normally deterministic (quarterly assessments, periodic guard patrolling, etc.) rather than being oblivious and temporary [12]. In addition, as information systems become more sophisticated in size and structure, and the motivation and nature of attacks change, it is becoming more difficult at any moment to be certain whether resources are secure, hence allowing "stealthy" moves to be realistic. In the next sections, we propose an extension to the above model as an attempt for the defender to more efficiently counter such moves.

3 Test it before Flipping it

The original FlipIt game models types of strategies for a player to regain control of a resource (i.e. to move) based on some pre-defined or on-the-fly tactics,

which however possess some limitations. In particular, a player may waste many moves if they happen while it is still controlling the resource. This becomes more serious if its periodic movement is significantly faster than the opponent's. Even if a move really serves its purpose, i.e., to regain control, it may still be an "almost" waste. This happens, for example, when the opponent's move is immediately (but coincidently) after such a move, rendering it ineffective.

Rather than blindly moving, an interesting question is whether knowing the state of control would be more beneficial to a player. In terms of information security assessment, this can be represented by the question "are we compromised?". The intuition behind this addition is rather simple. Knowing the state of control would prevent a waste move while the resource is still at hand. Also, even though it may not prevent an "almost" waste, it may suggest a timely response to a lost of control. This, of course, depends on how regularly the knowledge of the control state is updated.

To model such situations, we introduce a new class of strategies to FlipIt, namely the *state checking* strategies. As opposed to the ability to move/flip, a player is now able to check the game state, and then move/flip if necessary. In particular, we consider a strategy class $S = \{S_\alpha | \alpha > 0\}$ such that, given a strategy $S_\alpha \in S$, with $\delta = 1/\alpha$, player i may:

- perform a periodic state checking with period δ and cost u_i, with the first check occurring at a uniformly random time phase, i.e., within $[0, \delta]$;
- if a state check indicates a loss of control, immediately perform a move/flip (at cost k_i) to regain its control.

In addition to the original game FlipIt(P, P), several games might be introduced given S, for example FlipIt(S, P), FlipIt$(S \cup P, P)$, and FlipIt$(S \cup P, S \cup P)$. To study such games, it is important to notice that in all cases, the expected control time for each player can be formulated in the same way as that in $\{P, P\}$, using only δ_0 (or α_0) and δ_1 (or α_1). Indeed, at a time t, if a player is occupying the resource, a blind move action and a check-then-move action would yield the same effect, i.e., allowing it to regain control. Likewise, while it is in control of the resource, neither of the moves would bring any change. As this happens independently of the opponent's strategy, the same expected control time can be used for any game with strategies restricted to S and P.

Since a player's utility depends only on its expected control time and the cost of moving and/or checking, it is also independent of the opponent's type of strategy. Indeed, player 0 with strategy P_{α_0} would for example have a benefit as mentioned in Section 2

$$\beta_0(\alpha_0, \alpha_1) = \begin{cases} 1 - \frac{\alpha_1}{2\alpha_0} - k_0\alpha_0 & \text{if } \alpha_0 \geq \alpha_1 \\ \frac{\alpha_0}{2\alpha_1} - k_0\alpha_0 & \text{if } \alpha_0 < \alpha_1 \end{cases}.$$

With a strategy S_{α_0}, the average state checking cost for player 0 is $\alpha_0 u_0$. For moving cost, since S_{α_0} is employed, no move is wasted, thus player 0's number of moves is at most player 1's number of moves, i.e., $\min(\alpha_0, \alpha_1)$. This allows

the construction of its utility to become

$$\beta_0(\alpha_0, \alpha_1) = \begin{cases} 1 - \frac{\alpha_1}{2\alpha_0} - u_0\alpha_0 - k_0\alpha_1 & \text{if } \alpha_0 \geq \alpha_1 \\ \frac{\alpha_0}{2\alpha_1} - u_0\alpha_0 - k_0\alpha_0 & \text{if } \alpha_0 < \alpha_1 \end{cases}. \quad (1)$$

Given this new type of strategies S, a natural approach is to compare between S and P, that is, in which situations one is preferred over the other. The following theorem provides such comparison based on the relation between the costs of moving and state checking[2].

Theorem 1. *In the game FlipIt($P \cup S$, $P \cup S$), if $u_i \leq k_i/4$, player i does not prefer periodic moving. Otherwise, when $u_i \geq k_i$ player i does not prefer state checking.*

Proof. This theorem can be proved as a special case of Theorem 5, when $p = 1$.

Corollary 1. *Consider the game FlipIt($P \cup S$, $P \cup S$) with $k_i/4 < u_i < k_i$. Player i prefers a state checking strategy if and only if $\alpha_{1-i} \leq \frac{2(\sqrt{k_i}-\sqrt{u_i})^2}{k_i^2}$.*

The above results point out that when the cost of checking is sufficiently low, i.e., at most a quarter of the moving cost, it is always worth performing a check-then-move strategy. Indeed, as a low checking cost suggests a frequent checking schedule, a player is more closely up-to-date with its state of control of the resource. This helps the player to improve its expected control time, while keeping the moving cost at a reasonable level by eliminating wasted moves. Conversely, it is also intuitively clear that when the cost of checking exceeds that of moving, it is unreasonable to perform checking-then-moving. Furthermore, Corollary 1 indicates that, when the two cost are comparable, the best response for the opponent playing too fast is to either simply move at every step or not play at all, because at every step it is likely that without state checking the player is aware of its loss of control of the resource.

In the realm of information security, many situations may suggest that state checking strategies indeed outperform their moving counterparts. Consider an information system as the resource; the defender's act of moving/flipping is often expensive, as it might involve resets and restores of the system. This becomes more serious for large organisations, or those that require uninterrupted, real-time system availability and reliability, such as e-commerce, large computing facilities. On the other hand, checking for successful take-over of the system might be significantly cheaper and thus can be performed frequently, using intrusion detection systems (IDSs), auditing schemes, logging, etc. In such cases, it is recommended that funds are allocated for more frequent auditing of the system security to maximise the organisational benefit from the information system.

In another aspect, we recall from [14] that the game FlipIt(P, P) has a Nash equilibrium. As this game behaves similarly to infinitely repeated games, the equilibrium indicates the stage to which the game would eventually converge

[2] all proofs can be found in the full version of this paper at `http://eprint.iacr.org`.

if both players kept adjusting their actions upon realisation of the opponent's action. In the game FlipIt(S, S) however, such stable stage does not exist, as we show in Theorem 2. The intuition behind is reasonably simple. We notice that the total moving cost, i.e., $k_i \min(\alpha_i, \alpha_{1-i})$ for each player does not just depend on that player's rate, but also on its opponent. Thus, if a player keeps increasing its rate until it is faster than the opponent's, then its total moving cost will stop rising. This in turn results on a better chance for that player to come across a rate (possibly faster than the opponent) yielding higher benefit. This fact emphasises that when such situation occurs, the players' strategies are unstable, and it is best for a player to always monitor its opponent's state checking frequency and adjust its accordingly. In real life, this lack of stability suggests that the defender must keep consulting the statistics on how often attacks occur and adapt its strategy accordingly.

Theorem 2. *The Game FlipIt(S, S) has no pure strategy Nash equilibrium.*

4 Hardening Control over Time

Besides reactive measures such as state checking and moving, a proactive concern is on how to prevent losses of control from happening. In many cases this is more desirable because it is possible that consequences from attacks might have been overlooked, and thus it is better that attacks are prevented given the current realisation of potential losses. In the context of FlipIt, it may mean, for example, preventing a player from participating in the game, or to stop it after the game has run for some time. Following the analysis of the original FlipIt game, as well as those involving state checking strategies, it is not difficult to see that in order for a player to stop its opponent from participating in the game, it needs to play quick enough. Based on the best response functions for periodic moving and periodic state checking players, the rate limit above which player i should play so as to prevent its opponent from engaging on the game is

$$\alpha_i^{threshold} = \max\left(\frac{1}{2k_{1-i}}, \frac{k_{1-i} + u_{1-i} - \sqrt{u_{1-i}(2k_{1-i} + u_{1-i})}}{k_{1-i}^2}\right).$$

While this is desirable, it is sometimes infeasible to play fast enough if the state checking cost is high. A different preventive approach for a player is to somehow make it increasingly more difficult for its opponent to take over the resource over time. When the level of difficulty reaches some threshold, its opponent will automatically cease playing, and thus resulting in a long-term benefit for the player. In FlipIt type of games, this can be modelled by having a player spending an additional *periodic hardening cost* h_i every time it regains control, so that the opponent would have to spend more and more whenever trying to take over the resource. This cost could feature, for example, some penetration testing process that results in vulnerabilities being patched, similar to that modelled in [3]. It modifies the net utility of player i who performs state checking

with hardening as follows, with $m_i(t)$ being the number of state checks occurred prior to t:

$$B_i(t) = G_i(t) - (k_i + h_i)n_i(t) - u_i m_i(t).$$

In this section we aim to study how the defender selects its strategy based on the observed attacker's period. For the game analysis, we note that the utility of a player is represented by its average benefit since the game is infinite. The game in this section is however finite, and thus it is more reasonable to represent player i's utility as its net benefit over the whole game, i.e., $B_i(t_{end})$ where t_{end} is the moment in which the game ends. Assume that given a hardening cost h_0, the game ends after s state-changing attacks (i.e. flipping the state from the defender to the attacker). Since it is not difficult to see that such an attack occurs for every $\max(\delta_0, \delta_1)$ period, we may assume for simplicity that

$$t_{end} = s \cdot \max(\delta_0, \delta_1).$$

To analyse this game, we first model the utility function for each player. This can be done with two cases similar to the previous games.

- $\alpha_0 \geq \alpha_1$: similar to other periodic FlipIt games, the expected control time for the defender (i.e. player 0), is $(1 - r/2)\delta_1$ per δ_1, for $r = \alpha_1/\alpha_0$. We thus have the defender's utility as:

$$\begin{aligned} B_0(s\delta_1) &= (1 - \frac{r}{2})t - (k_0 + h_0)n_0(t) - u_0 m_0(t) \\ &= (1 - \frac{\delta_0}{2\delta_1})s\delta_1 - (k_0 + h_0)s - u_0 s\frac{\delta_1}{\delta_0} \\ &= s\left[\left(1 - \frac{\delta_0}{2\delta_1}\right)\delta_1 - k_0 - h_0 - u_0\frac{\delta_1}{\delta_0}\right]. \end{aligned}$$

  However, since different choices of h_0 yield different end times $t_{end} = s\delta_1$, it would be unreasonable to consider utility as the net benefit only until t_{end}. Indeed, consider h_0 and h_0' that yield ending time t_{end} and t'_{end} with net benefit B_0 and B_0', respectively, such that $t_{end} < t'_{end}$ and $B_0 < B_0'$. Even though $B_0 < B_0'$, this does not mean that the defender would prefer h_0' over h_0 since within the interval $[0, t'_{end}]$, the defender's net benefit would be $B_0 + (t'_{end} - t_{end})$, which might still be greater than B_0'.
  To resolve this issue, consider two choices of hardening costs h_0 and h_0' yielding different attack times s and s', with $s' > s$. The defender's net benefit within $[0, s'\delta_1]$ in these cases are respectively

$$B_0^* = B_0 + (s'\delta_1 - s\delta_1) = s\left[\left(1 - \frac{\delta_0}{2\delta_1}\right)\delta_1 - k_0 - h_0 - u_0\frac{\delta_1}{\delta_0}\right] + \delta_1(s' - s)$$

$$\text{and} \quad B_0' = s'\left[\left(1 - \frac{\delta_0}{2\delta_1}\right)\delta_1 - k_0 - h_0' - u_0\frac{\delta_1}{\delta_0}\right].$$

  By subtracting the latter to the former we get:

$$B_0' - B_0^* = s\left[\frac{\delta_0}{2\delta_1}\delta_1 + k_0 + h_0 + u_0\frac{\delta_1}{\delta_0}\right] - s'\left[\frac{\delta_0}{2\delta_1}\delta_1 + k_0 + h_0' + u_0\frac{\delta_1}{\delta_0}\right].$$

This implies that h_0' is preferred over h_0 if and only if

$$s\left[\frac{\delta_0}{2\delta_1}\delta_1 + k_0 + h_0 + u_0\frac{\delta_1}{\delta_0}\right] \geq s'\left[\frac{\delta_0}{2\delta_1}\delta_1 + k_0 + h_0' + u_0\frac{\delta_1}{\delta_0}\right].$$

As a result, we may effectively represent the defender's utility function in the following form:

$$U_0(\delta_0, h_0) = -s\left[\frac{\delta_0}{2\delta_1}\delta_1 + k_0 + h_0 + u_0\frac{\delta_1}{\delta_0}\right], \tag{2}$$

where the defender's action is a pair $(\delta_0, h_0) \in H_0$ implying the chosen state checking frequency (period) and hardening cost.

- $\alpha_0 \leq \alpha_1$: let $r = \delta_1/\delta_0 = \alpha_0/\alpha_1$. With similar reasoning as in the previous case, together with $n_0(t) = n_1(t) = s$ (due to alternating control) and $n_0(t) = m_0(t)$ (since $\alpha_0 \leq \alpha_1$) we have that the defender's net benefit is

$$\begin{aligned} B_0(s\delta_0) &= \frac{r}{2}t - (k_0 + h_0)n_0(t) - u_0 m_0(t) \\ &= s\left(\frac{\delta_1}{2\delta_0}\delta_0 - k_0 - h_0 - u_0\right). \end{aligned}$$

This leads to the defender's actual utility function as:

$$U_0(\delta_0, h_0) = -s\left[\left(1 - \frac{\delta_1}{2\delta_0}\right)\delta_0 + k_0 + h_0 + u_0\right]. \tag{3}$$

To complete the defender's utility function, it is important to compute s, the number of attacks, from the hardening cost h_0 and the original attack cost k_1. This can be generally modelled with a function f, such that at the s-th attack, the attack cost becomes $f_{h_0}^{s-1}(k_1)$, where $f_{h_0}(k_1) = f(k_1, h_0)$ gives the new cost of an attack due to h_0. The attacks stop at the $(s+1)$-th attempt if the cost involved is greater than the attacker's expected control, i.e.,

$$u_1 + f^s(k_1, h_0) \geq \max\left(\frac{\delta_0}{2}, \delta_0 - \frac{\delta_1}{2}\right).$$

In reality, the structure of f strongly depends on how control of the resource can be hardened. For example, if the resource contains a large number of identical and independent subsystems, so that the control becomes more secure as more subsystems are hardened, then one may model f as

$$f(k_1, h_0) = k_1 + \lambda h_0, \tag{4}$$

with $\lambda \geq 0$ signifies how effective the hardening process is. Another method is to follow an idea similar to that from Gordon and Loeb [7], in which the new cost of attack increases as more is spent on hardening the control. However, such increase should not be linear as in (4), but at a decreasing rate. Also, [4] and [3] suggest a weakest-link model in which attack cost increases linearly step by

step. Based on these results, we devise another reasonable construction for f as follows:

$$f(k_1, h_0) = k_1 + \frac{\mu h_0}{h_0 + \lambda}, \tag{5}$$

where $\mu \geq 0$ is the least upperbound on the increase of attack cost, and $\lambda > 0$ represents the effectiveness of the hardening process, so that it is more effective when λ is small. It is not difficult to see that since $f'(h_0) > 0$ and $f''(h_0) < 0$, the attack cost increases with the hardening cost, but at a decreasing rate, thus agreeing with Gordon and Loeb's model. Also, with the same hardening cost, the attack cost is raised by the same amount after each attack. Define $H = \{(\alpha, h) | \alpha > 0, h \geq 0\}$ to be a set of *periodic state checking with hardening* strategies as described above, we study in the following theorems recommendations for the defender in response to periodic attacks.

Theorem 3. *Consider the game FlipIt(H, $S \cup P$) with $f(k_1, h_0) = k_1 + \lambda h_0$. The defender's best response is $h_0 = B/\lambda$, where B is the attacker's expected utility for the first attack. The game ends after one successful attack.*

Theorem 4. *Consider the game FlipIt(H, $S \cup P$) with $f(k_1, h_0) = k_1 + \frac{\mu h_0}{h_0 + \lambda}$. Let $B > 0$ be the attacker's expected utility for the first attack, and let $n_a = B/\mu$. Let $L > 0$ be the defender's expected loss[3] per attack excluding h_0. Then*

- *for any hardening cost h_0, at least $\lceil n_a \rceil$ attacks occur before the game ends.*
- *the optimal hardening cost is*

$$h_0 = \frac{\lambda n_a}{s - n_a} \text{ where } s = \left\lceil n_a + \frac{1}{2}\left(\frac{\sqrt{L + 4n_a^2\lambda}}{\sqrt{L}} - 1\right)\right\rceil \tag{6}$$

is the corresponding number of attacks.

The above theorems stress a need for appropriate decision over the investment for hardening the resource control. In terms of information security, hardening may mean, for example, system patching, penetration testing, adding security layers, etc. However, an improvement in security does not necessarily imply a better return on security investment, as one can infer from Theorem 4. This happens when security does not just improve with the hardening cost, but depends on other factors, such as information. For example, a system may become more secure not via deployment of new measures, but rather because it gets fixed after suffering more and more attacks. While this idea is captured in (5), Theorem 4 suggests that the defender should spend enough to, for example, sufficiently patch the vulnerability, so that the attack cost would be raised by an amount close to μ. Any additional expense becomes less effective as the increase is bounded by μ. In contrast, situations modelled by (4) represent security that can be strengthened with little information. A common example is when an attack occurs against a device in a homogeneous network. In this case, it is always better to patch all devices, whether they have been compromised or not.

[3] This loss includes the attacker's occupation of the resource and the cost spent on protecting the resource.

5 Dealing with Complex Systems

In this section, we study a different extension to the model in Section 3 to capture situations in which the control of a resource might be difficult to measure, and that state checking might be inaccurate. This disproves an inherent but hidden assumption made in previous models, that with a cost u_i, player i can always determine who is in control of the resource. Again, it addresses another important issue with organisational information security by exacerbating the question "are we compromised?" by "how certain are we that we are compromised?". An answer to such question reflects not just how often security should be assessed, but also how the assessment should be done.

We extend the previous state-checking model with a probability p that the state check succeeds in determining a loss of control, applied to the defender only. The reason for such bias is obvious: while the defender must examine every component of its system as a mean of state checking, the attacker only needs to consider what it has previously compromised, which normally happens with certainty. To simplify our modelling, we explicitly make two assumptions as follows.

A1. There exists no false positive in state checking, i.e., no false alarm on attack exists.

A2. Once a false negative occurs, it will persist until the attacker's next interaction with the resource, i.e., either via a state check, or a move/flip.

Based on these assumptions, we may reformulate the defender's utility functions from what is given in (1), with the help of Lemma 1. It is important to notice that while the average state checking cost remains the same, the average flipping/moving cost lessened by a factor of p, since only a p-fraction of losses in control are followed by a flip/move. These yield the following utility function

$$\beta_0(\alpha_0, \alpha_1) = \begin{cases} p\left(1 - \frac{\alpha_1}{2\alpha_0}\right) - u_0\alpha_0 - pk_0\alpha_1 & \text{if } \alpha_0 \geq \alpha_1 \\ p\left(\frac{\alpha_0}{2\alpha_1}\right) - u_0\alpha_0 - pk_0\alpha_0 & \text{if } \alpha_0 < \alpha_1 \end{cases} \tag{7}$$

Lemma 1. *Consider the FlipIt games in which the defender plays a periodic state checking strategy. Then the defender's average control rate is $p\gamma$, where p is the success probability to detect an attack, and γ is the defender's average control rate when every state check occurs with certainty, i.e., when $p = 1$.*

Similar to the its predecessor, with this model we are also interested in the conditions under which state checking is preferred to mere flipping, and vice versa. This concern is reflected in Theorem 5, which generalises the result given in Theorem 1, and thus emphasises a preference for strategies involving inexpensive state checking, i.e., equal to at most a $p/4$-fraction of the flipping cost. The subtle threshold for the attack rate α_1 in (8) explains the fact that if the attacker infrequently interacts with the resource, then by assumption **A2** it is difficult to detect an attack, and thus periodic flipping is more desirable.

Theorem 5. *Consider the game FlipIt($P \cup S$, $P \cup S$) in which there is a probability p that the defender can detect a take-over attack with the state-checking action. The defender does not prefer periodic moving if*

$$u_0 \leq \frac{k_0 p}{4} \quad \text{and} \quad \alpha_1 \geq \frac{1}{2k_0} \min\left(1, \left[\frac{2(1-p)}{p}\right]^2\right). \tag{8}$$

Corollary 2. *Consider the game FlipIt($P \cup S$, $P \cup S$) in which there is a probability p that the defender can detect an attack. Let $\overline{\alpha_1}$ (resp. $\overline{\alpha_1}^*$) be the minimum value for the attacker's move rate α_1 to drop a periodic-moving (resp. state-checking) defender from the game. Then, $\overline{\alpha_1}^* \geq \overline{\alpha_1}$ if and only if $u_0 \leq k_0 p/4$.*

The need for $u_0 \leq k_0/4$ is further strengthened by Corollary 2 which addresses the situation when the attacker plays too fast, e.g., $\alpha_1 > 1/(2k_0)$, and periodic moving cannot afford for positive payoff, leading to the system being indefensible [14]. This issue becomes more realistic when the attacker is given chances to perform state checking, since in the information security realm, the attacker's state checking can be inexpensive, e.g., reconnecting to backdoors, re-logging in with stolen passwords, etc. In this case, periodic state checking is more robust as they survive higher attack rates.

Another intrinsic part of Theorem 5 is its implication over what is the right cost for state checking. Indeed, flipping in security often involves procedures with high certainty (system reset, backup restores, failovers, etc.), hence their costs are normally determined rather than decided. In contrast, an organisation may choose to invest arbitrarily in administering its security, for example through guard patrolling, antivirus software, firewalls, etc., subject to how much it desires the situation to be in control. While the goal is to satisfy the condition $u_0 \leq pk_0/4$, it is hindered by an inherent constraint that p typically decreases/increases with u_0, that is, less efforts for state checking yields less certainty on its effectiveness.

We study this issue by modelling the connection between u_0 and p, along with an environment parameter $v > 0$ specifying how effectively the amount u_0 might be spent. For example, this parameter may deteriorate as the resource becomes increasingly more sophisticated. On the other hand, it may increase with the skills of the team performing state checking. We can model p as the function of u_0, parameterised by v in the following way

$$p_v(u_0) = 1 - \frac{1}{vu_0 + 1}. \tag{9}$$

It is not difficult to see that, by modelling the probability of successful state checking as in (9), the value $1/v$ represents the cost required for detection of attacks to succeed with a fair coin-flipping chance, i.e., 50%. Note that this does not mean state checking with cost $u_0 \leq 1/v$ can be replaced by "coin-flipping detection" of attacks, as it may violate assumption **A1** to create many false positives, and hence waste moves would become a credible threat to the net utility. We now analyse the threshold under which the cost for state checking suggests it to overpower merely periodic flipping strategies.

Corollary 3. *Consider the game FlipIt($P \cup S$, $P \cup S$) in which there is a probability p that the defender can detect an attack, with p satisfying* (9)*. Then, if* $u_0 \leq k_0/4 - 1/v$ *and* $\alpha_1 \geq \frac{1}{2k_0} \min\left[1, \frac{4}{(u_0 v)^2}\right]$*, it is better for the defender to perform periodic state checking.*

From the threshold for state checking cost given in Corollary 3, we may also evaluate whether state checking is at all justifiable given specific characteristics of the environments. Indeed, if the productivity of information security is too low, i.e., $v \leq 4/k_0$, the use of state checking in most cases would not improve the overall utility, as too much cost is required to produce little benefit. This refers to situations when there is a mismatch between the scope of the resource being administered, and of the team performing administration, which means either the resource is too complex, or the administration is immature. In turn, such situations may apply to fast-growing organisations with slower catching-up with technology as well as security evaluation. Another example is with small to medium-sized firms whose businesses strongly rely on information systems, as many of them would spend little research in foreseeing the nontrivial impact of low security administration to the net income.

In overall, Corollary 3 recommends firms not just about hiring an administration team with highest quality-price ratio, but also to spend their concerns on easing the administration of their resource. In reality, the latter can be accomplished in a variety of ways, such as removing redudant components, restructuring the system toward simplification, avoiding complicated dependencies using separation of duties, etc. Otherwise, even the most desirable administration team might still be insufficient for a positive return on investment.

6 Conclusion

In this paper we investigate the concern on the choices of long-term strategic security plans for protecting organisational assets. These choices are represented by questions such as "are we vulnerable?" and "are we compromised?" This concern has become increasingly more important for large businesses as well as governmental units in the era where attackers are advanced, and have the resources to be persistent.

To do so, we extend the FlipIt game between an attacker and a defender periodically taking over a resource from each other, with the tradeoff between the cost of taking over, and the duration of the control. In our model, in addition from taking over, we allow players to check who is controlling the resource. We compare between blind take-over strategies and those that involve "check first, then take over", and show a threshold for the checking cost, under which the latter tactic is preferred.

In further extensions, we study strategic plans on how organisations would rationally invest in security improvement to discourage attackers. Our analysis on specific models proposed suggests that there are cases in which a system must suffer from many attacks to become sufficiently secure to deter attackers.

In reality, this is because security breaches serve as valuable information for improving system security. In another aspect, we relax our hidden assumption so that state checking might be incorrect, and study not just the frequency of security assessment, but also how quality-price-ratio may even discourage assessment of security. Since our models mostly deal with the defender's utility, the lessons learned may apply to not just advanced persistent threats (APTs), but also a pool of non-persistent threats that occurs with known frequency, e.g., from a community of underground hackers.

References

1. R. Bejtlich. Testimony before the USCC Hearing on "Developments in China's Cyber and Nuclear Capabilities", March 26, 2012. `http://www.uscc.gov/hearings/2012hearings/written_testimonies/hr12_03_26.php`.
2. C. G. Billo. Cyber warfare: An analysis of the means and motivations of selected nation states. Technical report, Institute for Security Technology Studies at Darthmouth College, 2004.
3. R. Böhme and M. Félegyházi. Optimal information security investment with penetration testing. In *Proceedings of the First international conference on Decision and game theory for security*, GameSec 2010, pages 21–37, Berlin, Heidelberg, 2010. Springer-Verlag.
4. R. Böhme and T. Moore. The iterated weakest link: A model of adaptive security investment. In *Workshop on the Economics of Information Security (WEIS)*, 2009.
5. E. Chabrow. Identifying undetected breaches identifying undetected breaches: How data scientists analyze big data to spot vulnerabilities, April 20, 2012. `http://www.bankinfosecurity.co.uk/interviews/identifying-undetected-breaches-i-1542`.
6. A. Coviello. Open letter to RSA customers, March 17, 2011. `http://www.rsa.com/node.aspx?id=3872`.
7. L. A. Gordon and M. P. Loeb. The economics of information security investment. *ACM Transactions in Information System Security*, 5(4):438–457, November 2002.
8. H. Kunreuther and G. Heal. Interdependent Security. *Journal of Risk and Uncertainty*, 26(2):231–249, March 2007.
9. R. Langner. Stuxnet: Dissecting a cyber warfare weapon. *IEEE Security & Privacy*, 9(3):49–51, 2011.
10. M. H. Manshaei, Q. Zhu, T. Alpcan, T. Basar, and J. Hubaux. Game Theory Meets Network Security and Privacy. Technical report, EPFL, 2010.
11. National Institute of Standards and Technology. Technical Guide to Information Security Testing and Assessment. Special Publication 800–115, 2008.
12. National Institute of Standards and Technology. Recommended security controls for federal information systems and organizations. Special Publication 800–53, 2009.
13. M. Raya, R. Shokri, and J. Hubaux. On the tradeoff between trust and privacy in wireless ad hoc networks. In *Proceedings of the third ACM conference on Wireless network security*, WiSec '10, pages 75–80, New York, NY, USA, 2010. ACM.
14. M. van Dijk, A. Juels, A. Oprea, and R. L. Rivest. Flipit: The game of "stealthy takeover". Cryptology ePrint Archive, Report 2012/103, 2012.

15. J. Vijayan. Breach, undetected since '05, exposes data on Kingston customers, July 17, 2007. http://www.computerworld.com/s/article/9027220/Breach_undetected_since_05_exposes_data_on_Kingston_customers.